

Weekly Report

11th, Dec. - 17th, Dec. 2017

Done:

1. Learn more about the implementation of system. Since we have decided to use web-based, we need to determine the front-end, back-end, frameworks, database.
 - As we're building a prototype (the data to render), front-end framework is optional. But we have to consider future situations. D3 is a must; Deck.gl may be used in the future so we adopt React.js which is required by deck.gl.
 - We have geo-located queries, thus we adopt MongoDB, which support multiple kinds of geo-based queries, like contain\intersection. The query performance is also guaranteed.
 - Node.js and MongoDB work better, thus we use node.
 - In the future, we may have MLs, which python backend may have more advantages; python works better with MySQL. Since we focus on prototype, we don't consider it currently. Both MySQL and MongoDB support geo-based query very well.
2. Discussion on the data, the tasks:

Two buddies will do a part time job in this project. They may also focus on some topics of the same data, like uncertainties in the data, in the data fusion, in the prediction. I introduced all the data I searched from the Internet, and all the ideas I have thought of, based on the report proposed last week.

We review several systems online, and we may start with a enhancement of WIFIRE project by UCSD. We may add more small views which depict some statistics for users to refer; we add visual comparison, history exploration views.
3. One buddy has consulted with more expertise students in wildfire prediction area, and draw the conclusion that it is not suitable for us to do the prediction since we do not have so many resources; we may focus on other topics like: we construct a heterogenous information network, do recommendations for experts for better analysis. We will discuss it next time.
4. 一个使用场景的设想:

用户想要看某个地区, 首先选中它;

 - 我们展现一些历史指标(诸如天气情况, AQI, 草树易燃指数等), 既可以是近期的(前七天), 也可以是长期的(每年这个时候的);
 - 我们展现其历史上的火灾情况: 其范围内, 其他时候的火灾; 对应时候的天气情况;其造成的损失伤害统计量;
 - 上述查看时, 支持滑动条滚动查看; 支持属性选择\数据过滤; 支持时间或者空间 aggregation; 支持时空上的 focus+context; 支持存档
 - 可视比较, 既可以在地图上比较, 也可以在对应的天气或其他情况上做比较; 需要计算一些简单指标来显示差异, MongoDB 似乎对此有不错的支持.
 - 提供一些推荐功能, 但作为 prototype 我们可能只能根据一些简单指标比如报告文本相似度\地理位置相似性\伤害程度相似性\气象情况欧氏距离相似性进行推荐, 让用户知道相似情况、对应的处置处理情况等等.

To Do:

1. Get the data and fuse them.
2. I have asked Minfeng, Xumeng for some code on the rendering on map or framework usage/ project building, I will learn more from them. A toy system should be built.
3. Write a more detailed plan and assign tasks to my buddies.

Papers:

EuroVIS STAR: *Social Media Visual Analytics* 北大陈思明的一篇综述, 其本人在这一块做了非常多的优秀的工作, 本文概括的非常全面, 对于想要深入这一块的很有参考价值. 其对其研究方法做了综述、文章关系、文章中的可视分析 pipeline 都做了总结, 和 Ross 那边的综述一样非常完善. 网络(图)、时空信息、文本是社交媒体三大实体, 基于此, VA 领域提供了六大种任务: 可视监控、特征提取、事件检测、异常检测、预测分析和情况感知. 其提到一些现有局限性比如 geotag 的数据比较少, 有些东西会有人的 bias 等都在我们后面工作可能会遇到.

PVis 2015 *Can Twitter Really Save Your Life? A Case Study of Visual Social Media Analytics for Situation Awareness* 文章邀请了 29 名领域专家, 验证他们一个基于社交网络的情境感知系统做调研, 最终得出了他们的系统, 或者说可视分析在情境感知将会有用的结论. 其文章中对于实时的东西仍然只是一个模拟, 专家提出了不少意见值得我们思考的, 比如如何处理数据不确定性的问题, 比如能否依据一个人历史发帖来判定这个人的 twitter 有没有用(似乎有这样的文章), 处理其他多媒体信息比如视频图片等; 不确定性的问题我感觉好像在之前文章确实不多, 不管是社交媒体还是应急管理可视化.

VIS 2015 *Interactive Visual Discovering of Movement Patterns from Sparsely Sampled Geo-tagged Social Media Data* 文章一定程度上解决了不确定性的问题, 通过聚合聚类加以上下文信息提示的功能; 但若在应急管理方面可能这个不一定完全适用. 但其聚类不失为一个思路; 由于没有实际数据暂时不好断言. 而且在应急情况下, 我觉得 Geo-tag 数据比例会大于 3%(这是平均值), 我们后面若敲定做这一块还会继续思考.

安排表

内容	DDL	Milestone
数据抓取	12.20	尝试不同数据接口可靠性稳定性等
前后端学习与实现	月底	